

음주 판단율 개선을 위한 음주전후 음성신호 유효 프레임 편차 분포 특성에 관한 연구

이원희¹⁾, 배성근²⁾, 박상범³⁾, 배명진⁴⁾

A Study on the Characteristics of Valid Frame Error Distribution Before and After Drinking for Improvement of Drinking Judgment Rate

Won-Hee Lee¹⁾, Seong-Geon Bae²⁾, Sang-Bum Park³⁾, Myung-Jin Bae⁴⁾

요 약

본 논문에서는 음성 분석을 이용하여 음주 전후의 음성특징을 비교하면서 음주여부를 판단할 수 있는 파라미터 추출을 연구하였다. 기존에 제안된 방법은 음주 전후 유효프레임의 위치를 추출한 뒤 편차를 이용하여 음주 식별을 하였다. 술을 마시면 여러 가지 음성신호의 변화가 일어난다. 이러한 특징을 이용하여 유효 프레임은 음주 후가 음주 전보다 발생지속시간이 길다는 점과 발생율이 비교적 높다는 점을 이용한 방법이다. 발전된 방법으로 단순히 유효 프레임 편차의 수치비교 특징만을 이용하는 것이 아닌 각각 발생자의 특징을 고려하고 분류하여 최종적으로 판단율을 개선하기 위하여 분포를 이용하여 음주전후를 판단하고자 한다. 제안된 방법은 기존 음성신호뿐만 아닌 여러 환경에서 왜곡된 음성파형에서도 음주여부 판단이 강인하였다. 즉 음성신호가 왜곡 되더라도 파형의 특징과 달리 내재되어 있는 운율, 발성상태 같은 특징이 남아있고 데이터 집합수가 많을수록 좀 더 높은 음주 판단율 할 수 있는 결과를 얻을 수 있다. 이러한 방법을 이용하여 음성신호를 가지고 음주전후 음성신호를 적용하였을 때 여러 환경에서도 판단율이 강인한 결과를 얻게 되었다.

핵심어 : 음주, 취한 상태, 유효-프레임, 분포, 편차

Received(June 8, 2018), Review Result(June 17, 2018), Accepted(July 13, 2018), Published(August 31, 2018)

¹⁾(Student) Dept. Information and Telecommunication Engineering, SoongSil Univ., 369 Sangdo-Ro, Dong-jak-gu, Seoul, Korea
email: vbluelovev@ssu.ac.kr

²⁾(Professor) Div. Computer Media Information Engineering, Kangnam Univ., Gugal-dong, Giheung-gu, Yongin-si, Gyeonggi-do, Korea
email: sgbae@kangnam.ac.kr

³⁾(Student) Dept. Information and Telecommunication Engineering, SoongSil Univ., 369 Sangdo-Ro, Dong-jak-gu, Seoul, Korea
email: sbpark8510@naver.com

⁴⁾(Professor, Corresponding Author) Dept. Information and Telecommunication Engineering, SoongSil Univ., 369 Sangdo-Ro, Dong-jak-gu, Seoul, Korea
email: mjbae@ssu.ac.kr

Abstract

In this paper, we investigated the parameter extraction to determine whether to drink or not by comparing voice characteristics before and after drinking using voice analysis. The proposed method extracts the valid frame position before and after drinking and identifies the alcohol by using the deviation. Drinking leads to various changes in the voice signal. Using this feature, the effective frame is a method using the fact that the duration of vocalization is longer than that before drinking and the vocalization rate is relatively high. In order to improve the judgment rate by considering the features of each speaker instead of using only the numerical comparison feature of the effective frame deviation as an advanced method, it is tried to determine before and after drinking by using the distribution. The proposed method is robust to judge whether or not to drink alcoholic beverages even in distorted voice waveforms in various environments as well as existing voice signals. In other words, even if the voice signal is distorted, features such as the rhythm and vocalization state inherent to the characteristic of the waveform remain unlike the waveform characteristic, and the higher the number of data sets, the higher the judgment result of drinking can be obtained. When the speech signal was applied before and after drinking with the speech signal using this method, the judgment rate was robust even in various environments.

Keywords : Drinking, Intoxication, Valid-Frame, Distribution, Deviation

1. 서론

오늘날에는 음주문화와 술의 종류 등 다양하게 변화하고 있다. 음주를 단순히 즐기는 것이 아닌 맛과 분위기 등 부가적인 부분에서도 달라지고 있다. 이런 변화를 통해 술의 소비량은 매년 많이 증가하고 있다. 술을 마시게 되면 사고능력이 저하되어 자기 절제가 힘들어진다.

이로 인해 문제점도 발생하는데 그중에 하나는 음주 후에 운전이나 운항을 하는 것이다. 음주운전은 개인뿐만 아니라 다른 사람에게도 피해를 주기 때문에 살인운전이라고도 볼 수 있다. 이러한 피해를 줄이기 위해 음주단속을 시행하는데 음주 측정은 대부분 차가 다니는 차선, 선박 같은 경우 바다위에 세운 뒤 측정을 하고 있다. 음주측정은 음주측정기가 근거리에 있어야하며, 측정시행자가 항상 사고위험에 노출이 되어있고, 측정당사자는 음주측정 거부행사, 인권침해 같은 문제점이 존재한다. 그리고 철도, 선박, 항공기, 자동차 등 수송기관의 운행 중에는 근접거리에서 측정하기가 어려워 사후 적발해야 한다.

근거리가 아닌 원거리로 음성으로 음주를 측정을 하게 되면 선박이나 기차 같은 수송기관은 사후 적발이 아닌 사전 단속을 할 수 있으며, 인권침해나 측정 거부행사에 대해 해결책이 될 수 있다. 음성을 이용하여 음주측정을 하는 방법은 계속 연구되고 있다. 기존의 방법은 음주 여부를 판단하는 기준으로 피치나 포먼트 또는 에너지의 변화 가지고 측정한다. 또한 음주 시 비음화 현상으로 인해 음성신호의 주파수영역에서 고주파쪽 변화가 두드러지기 때문에 고주파 대역에 포커스가 맞춰서 진행되고 있다.[6-10]

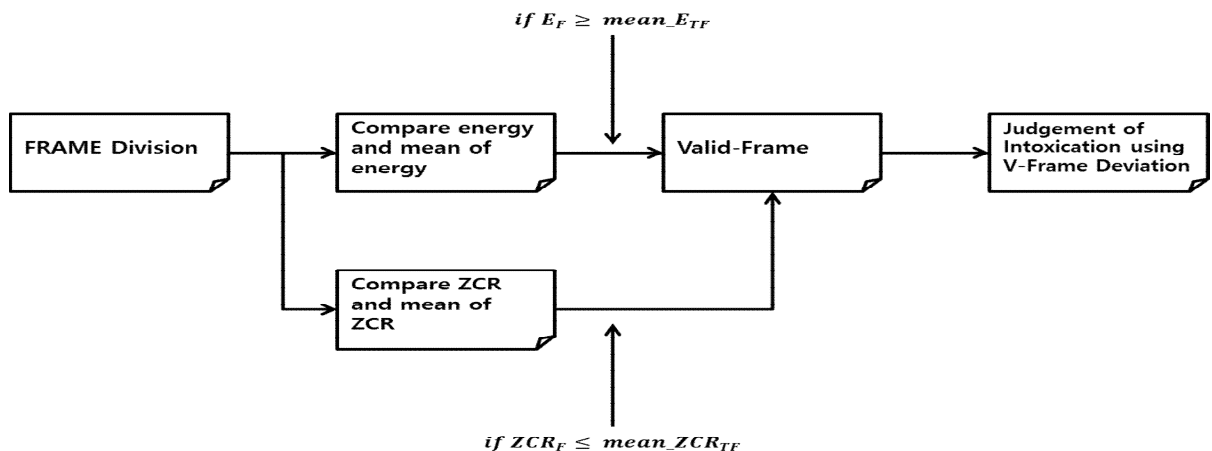
하지만 고주파 변화에 대한 음주판단 방법은 주변 환경 변화에 따라 특징이 저주파 영역에 비해 왜곡이 될 가능성이 크기 때문에 판단율이 저하가 되는 현상이 나타난다. 따라서 본 논문에서는 주변 환경 변화에 덜 민감한 음주 판단방법 중에 하나인 유효프레임 편차 측정 방법을 좀 더

디테일 하게 개선시켜 음주 판단율을 개선하고자 한다.

2장에서는 기존 유효프레임 편차 음주 판단 방법에 대해 알아보고 3장에서는 본 논문에서 제안하는 방법을 서술한다. 마지막으로 4장에서 결론을 맺는다.

2. 기존 유효프레임 편차 음주 판단 방법

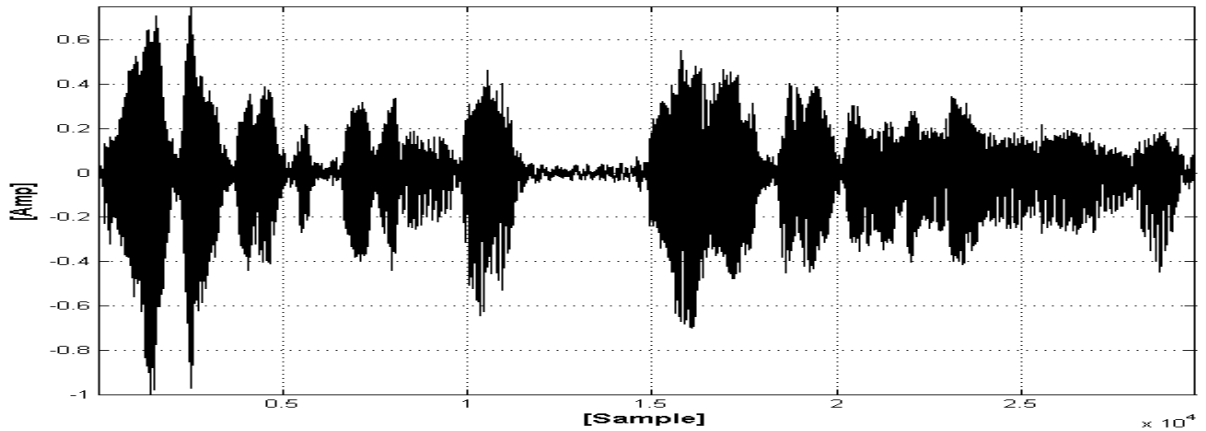
술을 마시면 여러 가지 음성변화가 일어나며 대표적으로 구강 점막이나 식도를 싸고 있는 표피층에 강한 자극을 주고 탈수현상을 일으키고 완전한 성대의 개·폐쇄가 일어나지 않는다.[3-4] 이러한 특징을 이용하여 사용되는 유효 프레임은 음주 후가 음주 전보다 발생지속시간이 길다는 점과 발생율이 비교적 높다는 점을 이용한 것이다. 따라서 음주 전이 음주 후보다 좀 더 유효한 프레임이 더 많다는 것을 의미한다고 볼 수 있다. 이러한 이유를 통해 유효프레임 거리 편차를 음주 전후의 음성에 적용하여 비교하는 것이 기존의 유효프레임 편차를 이용한 음주 검출방법이다. 그림 1은 기존의 유효프레임 편차를 이용한 음주 검출방법을 간단한 블록도로 나타낸 그림이다.



[그림 1] 기존 유효 프레임 편차를 이용한 음주 판단 블록도

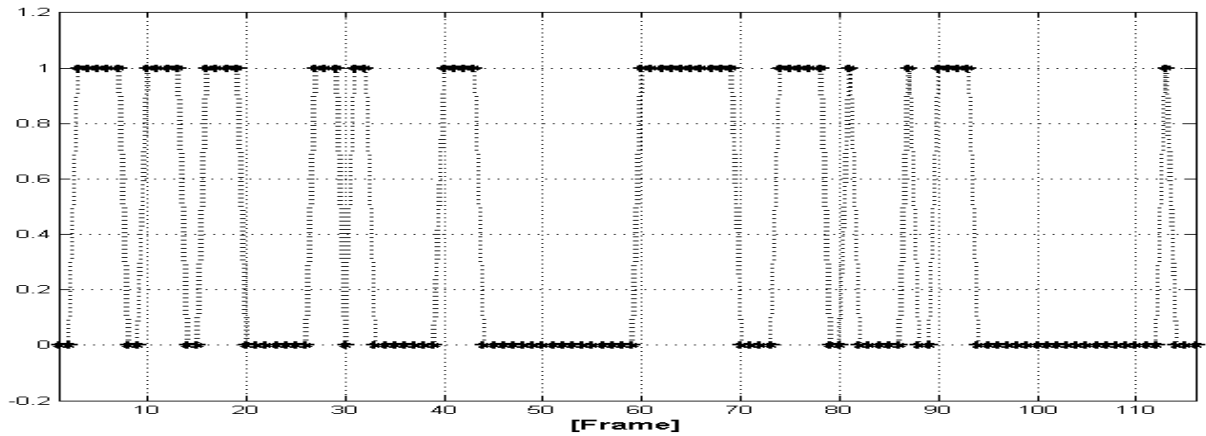
[figure 1] Intoxication judgement block diagram using existing Valid frame deviation

그림 1처럼 유효프레임의 기준 값을 이용하여 유효프레임을 추출하고 위치에 대한 표준편차 값을 이용하여 음주 전후 음성을 종속적으로 비교한다. 그림 2부터 그림 5은 음주 전후의 음성과정과 그림 1의 블록도를 이용한 유효프레임 추출된 결과 그래프이다. 그림 3과 그림 5의 Y축 결과 값이 1인 것은 유효프레임이며 0의 값은 유효프레임이 아닌 것으로 결과를 나타낸 것이다.



[그림 2] 음주 전 음성파형

[figure 2] Intoxication judgement block diagram using existing Valid frame deviation

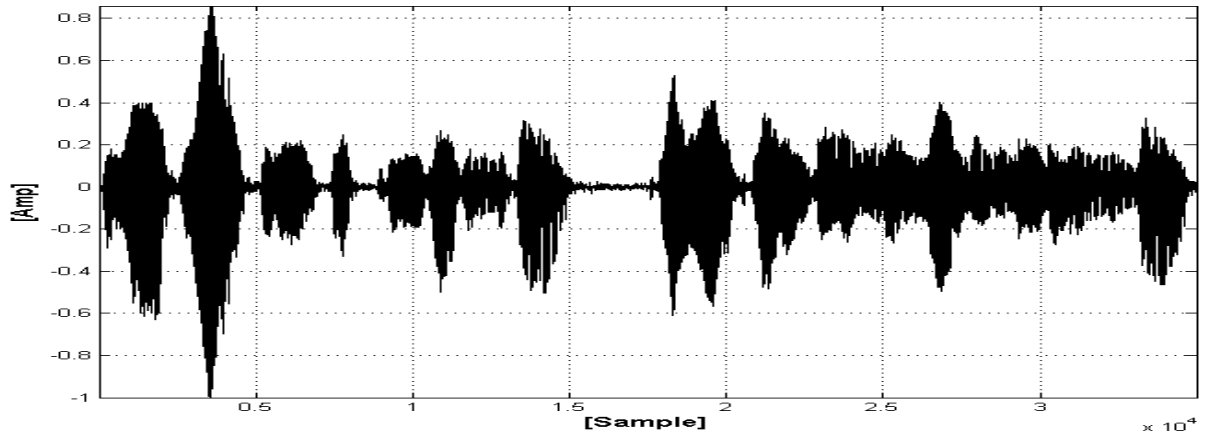


[그림 3] 음주 전 유효프레임 추출 결과

[figure 3] Intoxication judgement block diagram using existing Valid frame deviation

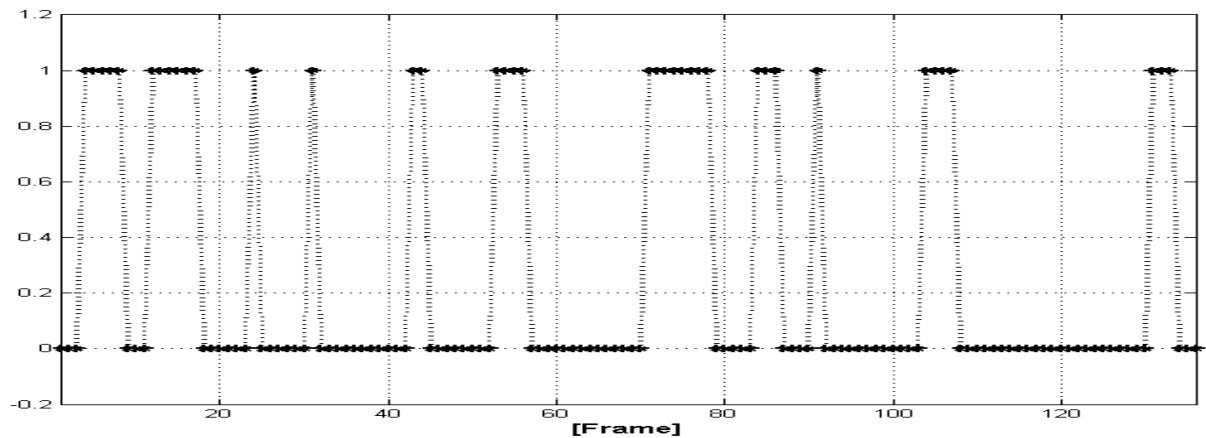
그림 3인 음주 전 유효프레임 추출 결과와 그림 5의 음주 후 유효프레임 추출 결과를 비교하면 음주 전과 비해 음주 후에는 유효프레임이 적게 추출이 되었으며 이로 인해 위치의 편차가 커지는 현상이 나타난다.

주파수 영역으로 음주전후를 비교할 시 클린 환경에서는 다른 영역보다는 정확한 음주판단이 가능하지만 주변 환경에 대한 영향에 따라 판단율이 편차가 다양하게 발생한다. 반대로 시간 영역에서는 주파수 영역에 비해 판단율의 정확도가 떨어지지만 주변 환경에서는 주파수 영역보다는 덜 민감하게 판단할 수 있는 장점이 있다.



[그림 4] 음주 후 음성파형

[figure 4] Intoxication judgement block diagram using existing Valid frame deviation



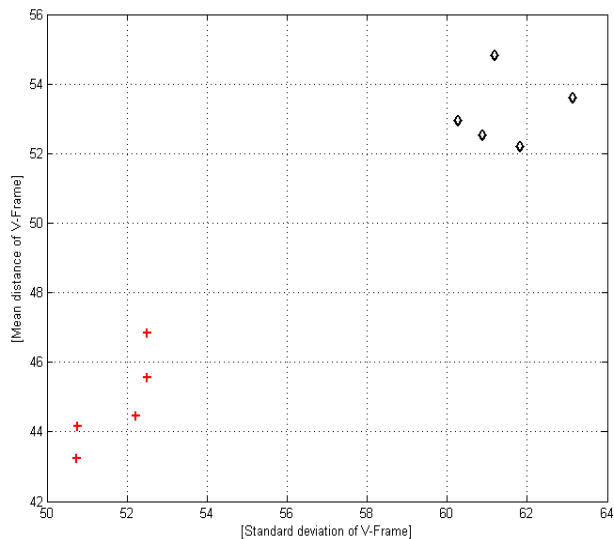
[그림 5] 음주 후 유효프레임 추출 결과

[figure 5] Intoxication judgement block diagram using existing Valid frame deviation

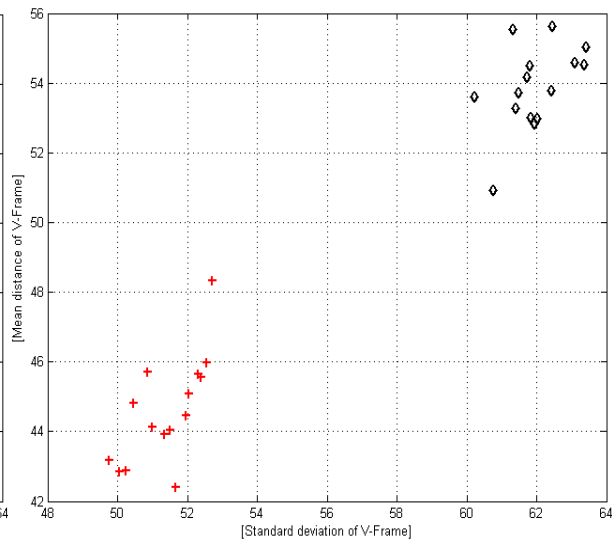
3. 제안하는 방법

본 논문에서는 기존에 제안된 유효 프레임을 이용한 음주판단 알고리즘을 통해 단순히 수치 값의 비교를 통한 음주 판단이 아닌 각각 발성자만의 음주전후 유효 프레임 파라미터 분포를 이용하여 최종적으로 음주 판단율을 개선하고자 한다. 그림 6부터 그림 9는 음주 전후 시 같은 사람이 똑같이 발성한 단문장 음원을 25개와 50개를 이용하여 유효 프레임 편차 데이터 특징을 나타낸 분포도이다.

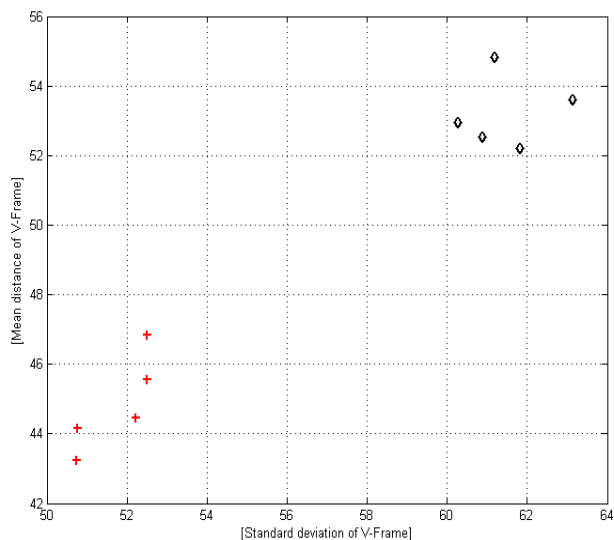
A Study on the Characteristics of Valid Frame Error Distribution Before and After Drinking for Improvement of Drinking Judgment Rate



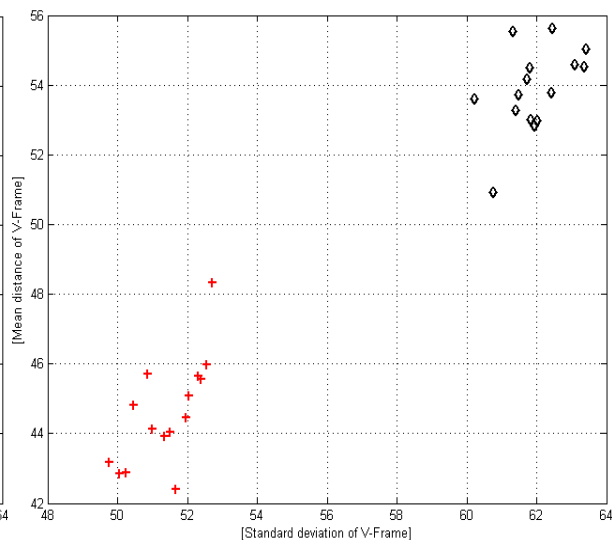
[그림 6] 같은 단문장 25개 음원(음주 전 랩실 환경) 가지고 나타낸 유효 프레임 데이터 집합 분포
 [figure 6] Distribution of valid frame data set showing with five speech sources of sentences of the same column



[그림 7] 같은 단문장 50개 음원(음주 전 랩실 환경) 가지고 나타낸 유효 프레임 데이터 집합 분포
 [figure 7] Distribution of valid frame data set showing with five speech sources of sentences of the same column



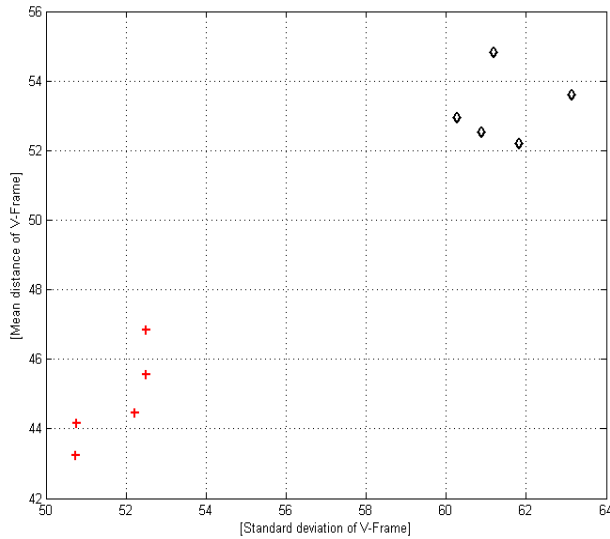
[그림 8] 같은 단문장 25개 음원(음주 후 랩실 환경) 가지고 나타낸 유효 프레임 데이터 집합 분포
 [figure 8] Distribution of valid frame data set showing with five speech sources of sentences of the same column



[그림 9] 같은 단문장 50개 음원(음주 후 랩실 환경) 가지고 나타낸 유효 프레임 데이터 집합 분포
 [figure 9] Distribution of valid frame data set showing with five speech sources of sentences of the same column

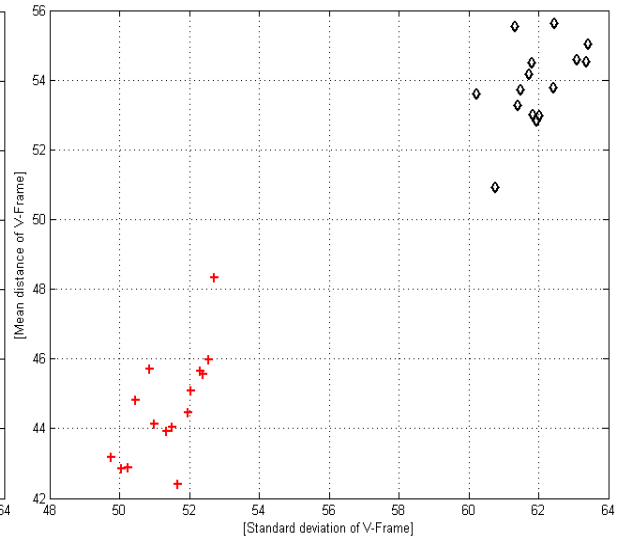
그림 6부터 그림 9의 빨강색은 음주 전 특징들의 데이터 집합이며, 검정색은 음주 후 특징을 나

타내는 데이터 집합이다. 음원의 개수 즉 데이터집합들이 늘어날수록 음주전후의 유효프레임 편차 분포 특성이 확실하게 구별되고 있다. 즉 종속적으로 비교방법 뿐만 아닌 독립적인 방법으로 음주 전후를 판단하기 위해서는 단순히 여러 사람들의 음성신호를 동시에 분석 및 파라미터를 추출한 뒤 어떠한 임계값 또는 기준 값을 정하는 것 이외에 각각 개인의 음성신호에 특성에 맞추어 판단할 수 있도록 유형에 대한 데이터 셋을 분류하고 구하는 것을 추가시키는 것이 최종적으로 판단을 개선시키는데 중요하다고 볼 수 있다. 그림 10부터 17은 각각 환경에 대한 유효프레임 특징에 대한 분포 그래프이다.



[그림 10] 같은 단문장 25개 음원(음주 전 15% 과부하 상태) 가지고 나타낸 유효 프레임 데이터 집합 분포

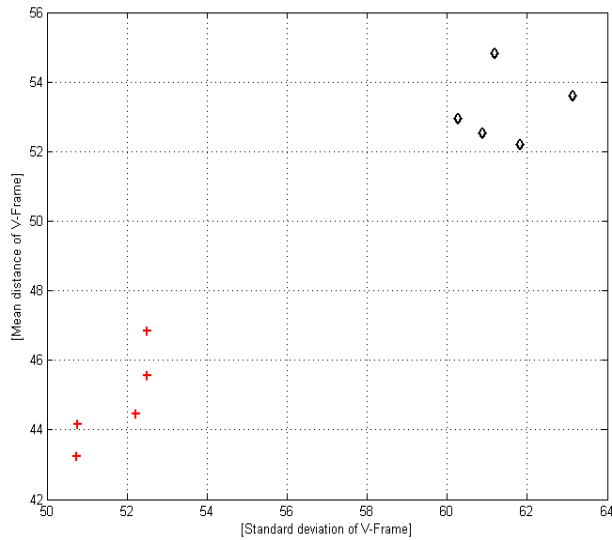
[figure 10] Distribution of valid frame data set showing with five speech sources of sentences of the same column



[그림 11] 같은 단문장 50개 음원(음주 전 15% 과부하 상태) 가지고 나타낸 유효 프레임 데이터 집합 분포

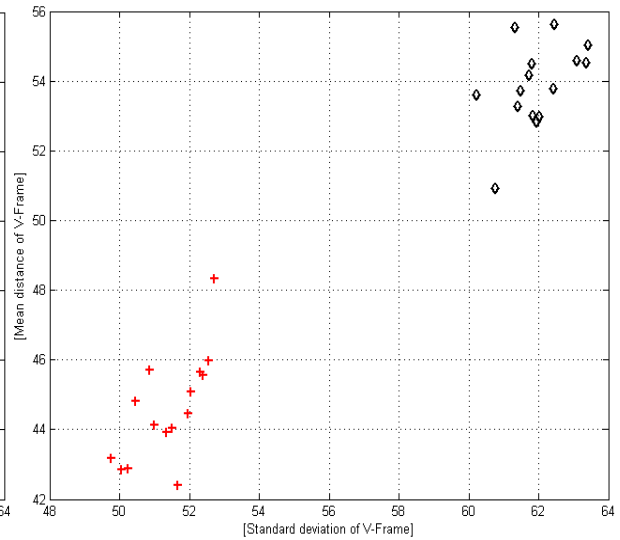
[figure 11] Distribution of valid frame data set showing with five speech sources of sentences of the same column

A Study on the Characteristics of Valid Frame Error Distribution Before and After Drinking for Improvement of Drinking Judgment Rate



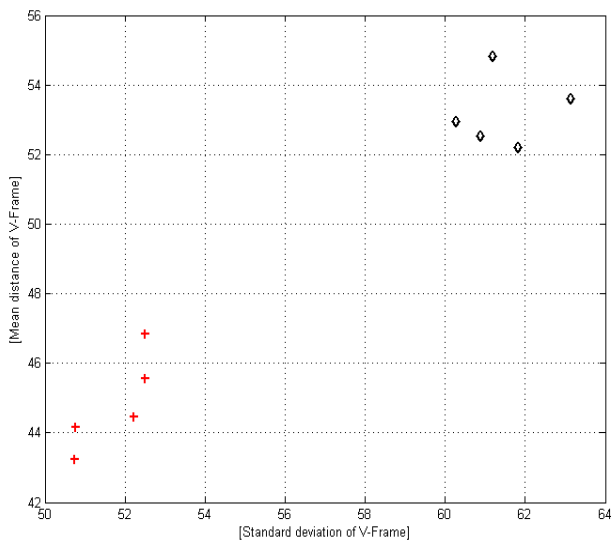
[그림 12] 같은 단문장 25개 음원(음주 후 15% 과부하 상태) 가지고 나타난 유효 프레임 데이터 집합 분포

[figure 12] Distribution of valid frame data set showing with five speech sources of sentences of the same column



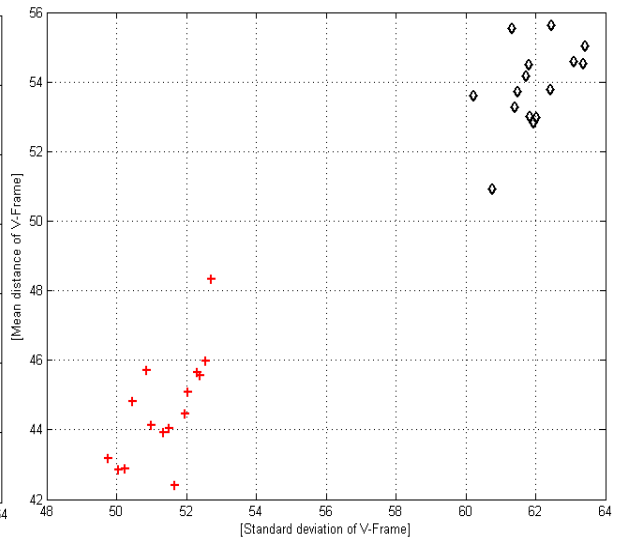
[그림 13] 같은 단문장 50개 음원(음주 후 15% 과부하 상태) 가지고 나타난 유효 프레임 데이터 집합 분포

[figure 13] Distribution of valid frame data set showing with five speech sources of sentences of the same column



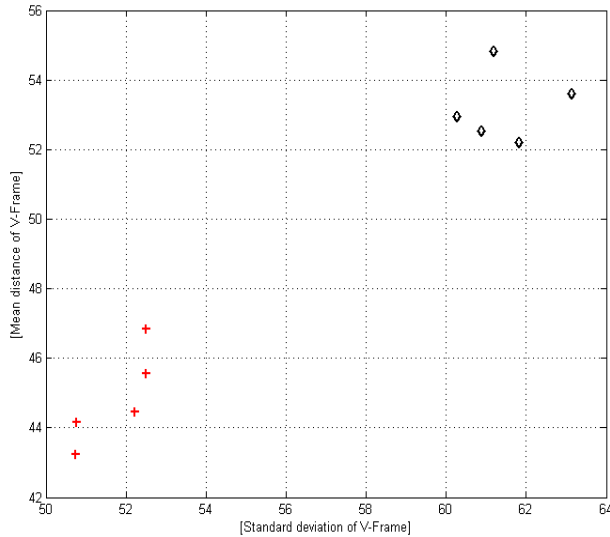
[그림 14] 같은 단문장 25개 음원(음주 전 엔진소리 환경) 가지고 나타난 유효 프레임 데이터 집합 분포

[figure 14] Distribution of valid frame data set showing with five speech sources of sentences of the same column



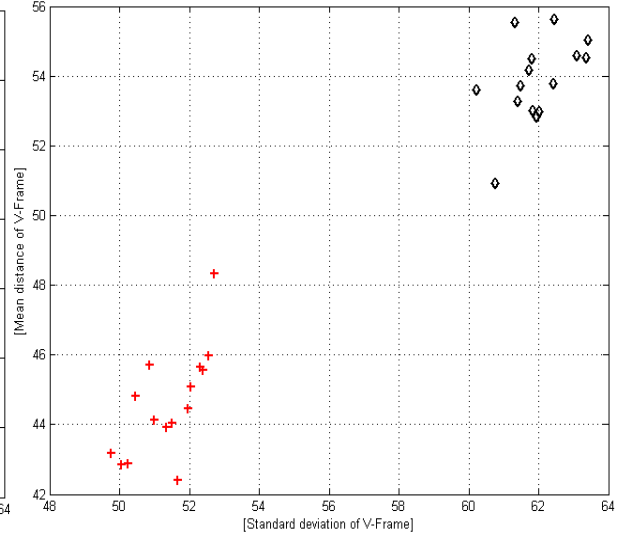
[그림 15] 같은 단문장 50개 음원(음주 전 엔진소리 환경) 가지고 나타난 유효 프레임 데이터 집합 분포

[figure 15] Distribution of valid frame data set showing with five speech sources of sentences of the same column



[그림 16] 같은 단문장 25개 음원(음주 후 엔진소리 환경) 가지고 나타낸 유효 프레임 데이터 집합 분포

[figure 16] Distribution of valid frame data set showing with five speech sources of sentences of the same column



[그림 17] 같은 단문장 50개 음원(음주 후 엔진소리 환경) 가지고 나타낸 유효 프레임 데이터 집합 분포

[figure 17] Distribution of valid frame data set showing with five speech sources of sentences of the same column

그림 10부터 그림 17을 보면 알 수 있듯이 각각에 대한 주변 환경이 다르더라도 음주전후에 대한 유효프레임 특징이 비교적 쉽게 분류 할 수 있을 정도로 분포가 되어있음을 알 수 있다. 또한 랩실환경과 각 환경을 비교해 보았을 때도 유효프레임 음주판단방법 특징에 따라 분류된 값들이나 위치가 크게 달라지지 않음을 알 수 있다.[1-2],[4] 결과적으로 제안된 방법은 결과 그래프에 따라 기존 음성신호뿐만 아닌 여러 환경에서 왜곡된 음성파형에서도 음주여부 판단이 강인하였다. 즉 음성신호가 왜곡 되더라도 파형의 특징과 달리 내재되어 있는 운율, 발성상태 같은 특징이 남아있고 데이터 집합수가 많을수록 좀 더 높은 음주판단을 할 수 있는 결과를 얻을 수 있다.

4. 결론 및 향후 연구방향

도로위에서 발생하는 음주사고는 날로 빈도와 규모가 커지고 있다. 이에 따라 국가차원에서 도로법률개정이나 단속시스템 등 여러 방면에서 음주사고를 미리 예방하고자 노력하고 있지만 비용과 시간 등 현실적인 한계로 인해 큰 어려움을 겪고 있다.

따라서 본 논문에서는 효율적으로 미리 음주여부를 판단할 수 있도록 음성신호의 음주 판단을 개선을 위한 유효 프레임 편차 분포를 연구하였다. 기존의 유효프레임 편차를 통한 음주판단방법은 음주 후가 음주 전보다 발성지속시간이 길다는 점과 발성율이 비교적 높다는 점을 이용한 것이다. 추가로 유효프레임을 이용한 음주판단방법의 판단율을 개선시키기 위해 개개인의 추출된 파라

미터의 분포에 대하여 분석을 진행하였다.

이러한 특징을 이용하여 개인 맞춤형 판단 뿐 만 아닌 여러 다양한 환경에서도 음주여부 판단율이 강인할 수 있다. 향후에는 이런 연구를 통해 다양한 시스템 환경에서도 유용하게 사용될 것으로 기대한다.

References

- [1] W. H. Lee, S. Y. Park, M. J. Bae, Editors. A Study on Comparing and Analysing the Character of Pitch-bandwidth of Drunk and non Drunk Voice. The Korean Institute of Communications and Information Science, (2014), pp.876-877
- [2] Geum-Ran Baek, Myung-Jin Bae, "Study on the Judgment of Intoxication State using Speech Analysis ". Vol.352 CCIS, 2012,, Springer-Verlag, pp. 277-282.
- [3] S. Marcieli Bellé, Silvia do Amaral Sartori, Angela Garcia Rossi. "Alcoholism: effects on the cochleo-vestibular apparatus". Rev Bras Otorrinolaringol. Vol73, No.1 (2007), pp116-222.
- [4] Won-Hee Lee, Myung-Sook Kim, and Myung-Jin Bae, Using valid-frame deviation to judgment of intoxication. Information : An International Interdisciplinary Journal, Vol. 18 (2015), pp4131-4136.
- [5] Seong-Geon Bae, Won-Hee Lee and Myung-Jin Bae, " A Study on Low Frequency Noise of Dehumidifier using Acoustic Characteristics," IJET Vol.8, No.1, pp.235-237.
- [6] Won-Hee Lee, Seong-Geon Bae and Myung-Jin Bae, "A Study on Improving the Overloaded Speech Waveform to Distinguish Alcohol Intoxication using Spectral Compensation," IJET Vol.7, No.5 (2015), pp.1957-1964.
- [7] Won-Hee Lee, Myung-Jin Bae, "Reducing Errors of Judgment of Intoxication in Overloaded Speech Signal," IJET Vol.8, No.1 (2016), pp.219-224.
- [8] Seong-Geon Bae, Myung-Sook Kim, and Myung-Jin Bae, Information : An International Interdisciplinary Journal, Using High Frequency Accentuation in Speech Signals as a New Parameter in Intoxication Judgment. (2014) Vol. 17, 6531-6537
- [9] Chang-Joong Jung, Seong-Geon Bae, Myung-Sook Kim and Myung-Jin Bae, IJMUE, "Speech Sobriety Test Based on Formant Energy Distribution, (2013) Vol.8, No.6, pp.209-216
- [10] Ji-Eon Jang, Journal of the Korea Entertainment Industry Association, The Relationship of Health Behavior, Oral Management Behaviors and Oral Health Status. (2015), Vol.9, No.3, pp.427-437